

(19)日本国特許庁 (JP)

## (12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-321177

(P2002-321177A)

(43)公開日 平成14年11月5日(2002.11.5)

(51)Int.Cl.<sup>7</sup>  
B 25 J 5/00  
A 63 H 3/33  
11/00  
B 25 J 13/00  
G 10 L 13/00

識別記号

F I  
B 25 J 5/00  
A 63 H 3/33  
11/00  
B 25 J 13/00  
G 10 L 3/00

テマコト\*(参考)  
C 2 C 1 5 0  
C 3 C 0 0 7  
Z 5 D 0 1 5  
Z 5 D 0 4 5  
5 3 7 B

審査請求 未請求 請求項の数29 O L (全 14 頁) 最終頁に続く

(21)出願番号 特願2001-124974(P2001-124974)

(22)出願日 平成13年4月23日(2001.4.23)

(71)出願人 000002185  
ソニー株式会社  
東京都品川区北品川6丁目7番35号  
(72)発明者 大久保 厚志  
東京都品川区北品川6丁目7番35号 ソニー株式会社内  
(72)発明者 ガブリエル コスター  
東京都品川区北品川6丁目7番35号 ソニー株式会社内  
(74)代理人 100101801  
弁理士 山田 英治 (外2名)

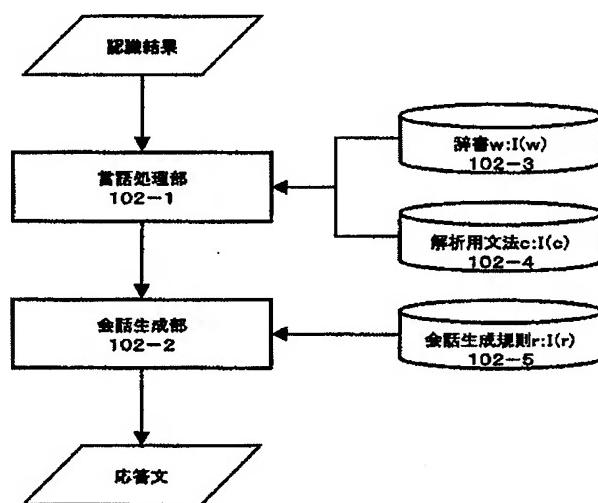
最終頁に続く

(54)【発明の名称】 脚式移動ロボット及びその制御方法

### (57)【要約】

【課題】 内部状態の成長の度合いに応じて対話機能を成長させる脚式移動ロボットを提供する。

【解決手段】 音声認識システムは、単語・構文認識機能を持ち、登録された単語や構文に知性レベルを対応付けておき、発話文と応答文の知性レベル、ロボットの成長度を比較することによって、認識された発話文に対する応答の発生確率を求めて、ロボットの成長過程に応じて対話内容を変化させていく。ロボットは、内部状態に成長に合わせて、最初は簡単な文しかしゃべらないし理解できないが、徐々に複雑な文に対して複雑な応答を行えるようになる。



**【特許請求の範囲】**

【請求項1】駆動可能な複数の関節角からなる脚式移動ロボットであって、  
ユーザの音声を認識する音声認識手段と、  
音声認識された発話文に対する前記脚式移動ロボットの知性レベルに相応した応答文を生成する会話処理手段と、  
生成された応答文を音声出力する音声合成手段と、を具備することを特徴とする脚式移動ロボット。

【請求項2】前記会話処理手段は、知性レベルに対応付けられた単語辞書、知性レベルに対応付けられた構文解析用文法、並びに知性レベルに対応付けられた会話生成規則に基づいて、知性レベルに相応した応答文を生成する、ことを特徴とする請求項1に記載の脚式移動ロボット。

【請求項3】前記会話処理手段は、認識された発話文を構成する単語及び／又は構文に対応付けられた知性レベルを基に、該発話文の知性レベルを求める、ことを特徴とする請求項1に記載の脚式移動ロボット。

【請求項4】前記会話処理手段は、認識された発話文の知性レベルと前記脚式移動ロボットの成長度の相対値に基づいて応答文を生成する、ことを特徴とする請求項1に記載の脚式移動ロボット。

【請求項5】前記会話処理手段は、認識された発話文の知性レベルと前記脚式移動ロボットの成長度の相対値に近い知性レベルの応答文を生成する、ことを特徴とする請求項1に記載の脚式移動ロボット。

【請求項6】前記会話処理手段は、認識された発話文の知性レベルと前記脚式移動ロボットの成長度の相対値に応じた各応答文の発生確率に従って応答文を選択する、ことを特徴とする請求項1に記載の脚式移動ロボット。

【請求項7】前記会話処理手段は、応答文の知性レベルと前記脚式移動ロボットの成長度に基づいて応答文を選択する、ことを特徴とする請求項1に記載の脚式移動ロボット。

【請求項8】前記会話処理手段は、前記脚式移動ロボットの成長度に最も近い知性レベルを持つ応答文を選択する、ことを特徴とする請求項1に記載の脚式移動ロボット。

【請求項9】前記会話処理手段は、前記脚式移動ロボットの成長度に応じた各応答文の発生確率に従って応答文を選択する、ことを特徴とする請求項1に記載の脚式移動ロボット。

【請求項10】駆動可能な複数の関節角からなる脚式移動ロボットの制御方法であって、  
ユーザの音声を認識する音声認識ステップと、  
音声認識された発話文に対する前記脚式移動ロボットの知性レベルに相応した応答文を生成する会話処理ステップと、  
生成された応答文を音声出力する音声合成ステップと、

を具備することを特徴とする脚式移動ロボットの制御方法。

【請求項11】前記会話処理ステップでは、知性レベルに対応付けられた単語辞書、知性レベルに対応付けられた構文解析用文法、並びに知性レベルに対応付けられた会話生成規則に基づいて、知性レベルに相応した応答文を生成する、ことを特徴とする請求項10に記載の脚式移動ロボットの制御方法。

【請求項12】前記会話処理ステップでは、認識された発話文を構成する単語及び／又は構文に対応付けられた知性レベルを基に、該発話文の知性レベルを求める、ことを特徴とする請求項10に記載の脚式移動ロボットの制御方法。

【請求項13】前記会話処理ステップでは、認識された発話文の知性レベルと前記脚式移動ロボットの成長度の相対値に基づいて応答文を生成する、ことを特徴とする請求項10に記載の脚式移動ロボットの制御方法。

【請求項14】前記会話処理ステップでは、認識された発話文の知性レベルと前記脚式移動ロボットの成長度の相対値に近い知性レベルの応答文を生成する、ことを特徴とする請求項10に記載の脚式移動ロボットの制御方法。

【請求項15】前記会話処理手段は、認識された発話文の知性レベルと前記脚式移動ロボットの成長度の相対値に応じた各応答文の発生確率に従って応答文を選択する、ことを特徴とする請求項10に記載の脚式移動ロボット。

【請求項16】前記会話処理ステップでは、応答文の知性レベルと前記脚式移動ロボットの成長度に基づいて応答文を選択する、ことを特徴とする請求項10に記載の脚式移動ロボットの制御方法。

【請求項17】前記会話処理ステップでは、前記脚式移動ロボットの成長度に最も近い知性レベルを持つ応答文を選択する、ことを特徴とする請求項10に記載の脚式移動ロボットの制御方法。

【請求項18】前記会話処理ステップでは、前記脚式移動ロボットの成長度に応じた各応答文の発生確率に従って応答文を選択する、ことを特徴とする請求項10に記載の脚式移動ロボットの制御方法。

【請求項19】ユーザからの発話文に対してキャラクタの成長度に応じた応答文を生成する対話制御方法であつて、

ユーザの音声を認識する音声認識ステップと、  
音声認識された発話文に対する前記キャラクタの知性レベルに相応した応答文を生成する会話処理ステップと、  
生成された応答文を音声出力する音声合成ステップと、  
を具備することを特徴とする対話制御方法。

【請求項20】前記会話処理ステップでは、知性レベルに対応付けられた単語辞書、知性レベルに対応付けられた構文解析用文法、並びに知性レベルに対応付けられた

会話生成規則に基づいて、知性レベルに相応した応答文を生成する、ことを特徴とする請求項19に記載の対話制御方法。

【請求項21】前記会話処理ステップでは、認識された発話文を構成する単語及び／又は構文に対応付けられた知性レベルを基に、該発話文の知性レベルを求める、ことを特徴とする請求項19に記載の対話制御方法。

【請求項22】前記会話処理ステップでは、認識された発話文の知性レベルと前記キャラクタの成長度の相対値に基づいて応答文を生成する、ことを特徴とする請求項19に記載の対話制御方法。

【請求項23】前記会話処理ステップでは、認識された発話文の知性レベルと前記キャラクタの成長度の相対値に近い知性レベルの応答文を生成する、ことを特徴とする請求項19に記載の対話制御方法。

【請求項24】前記会話処理手段は、認識された発話文の知性レベルと前記キャラクタの成長度の相対値に応じた各応答文の発生確率に従って応答文を選択する、ことを特徴とする請求項19に記載の対話制御方法。

【請求項25】前記会話処理ステップでは、応答文の知性レベルと前記キャラクタの成長度に基づいて応答文を選択する、ことを特徴とする請求項19に記載の対話制御方法。

【請求項26】前記会話処理ステップでは、前記キャラクタの成長度に最も近い知性レベルを持つ応答文を選択する、ことを特徴とする請求項19に記載の対話制御方法。

【請求項27】前記会話処理ステップでは、前記脚式移動ロボットの成長度に応じた各応答文の発生確率に従つて応答文を選択する、ことを特徴とする請求項19に記載の対話制御方法。

【請求項28】駆動可能な複数の関節角からなる脚式移動ロボットがユーザと対話するための処理をコンピュータ・システム上で実行するように記述されたコンピュータ・ソフトウェアをコンピュータ可読形式で物理的に格納した記憶媒体であって、前記コンピュータ・ソフトウェアは、

ユーザの音声を認識する音声認識ステップと、音声認識された発話文に対する前記脚式移動ロボットの知性レベルに相応した応答文を生成する会話処理ステップと、生成された応答文を音声出力する音声合成ステップと、を具備することを特徴とする記憶媒体。

【請求項29】ユーザからの発話文に対してキャラクタの成長度に応じた応答文を生成する対話処理をコンピュータ・システム上で実行するように記述されたコンピュータ・ソフトウェアをコンピュータ可読形式で物理的に格納した記憶媒体であって、前記コンピュータ・ソフトウェアは、

ユーザの音声を認識する音声認識ステップと、

音声認識された発話文に対する前記キャラクタの知性レベルに相応した応答文を生成する会話処理ステップと、生成された応答文を音声出力する音声合成ステップと、を具備することを特徴とする記憶媒体。

#### 【発明の詳細な説明】

##### 【0001】

【発明の属する技術分野】本発明は、駆動可能な複数の関節角からなる脚式移動ロボット及びその制御方法に係り、特に、ユーザとの対話機能を持つ脚式移動ロボット及びその制御方法に関する。

【0002】更に詳しくは、本発明は、内部状態の変化に応じた対話をを行う脚式移動ロボット及びその制御方法に係り、特に、内部状態の成長の度合いに応じて使用する語彙や応答のパターンなどを変化させて対話機能を成長させる脚式移動ロボット及びその制御方法に関する。

##### 【0003】

【従来の技術】従来、ユーザからの指令や周囲の環境に応じて動作する歩行型のロボット装置、多関節ロボット、あるいはコンピュータ・グラフィックス(CG)で動作するキャラクタを用いたアニメーションなどに関して、数多の提案がなされ、開発されている。このようなロボット装置あるいはアニメーション(以下では、これらをまとめて「ロボット装置等」とも呼ぶ)は、ユーザからの指令に応答して一連の動作を行う。

【0004】例えば、イヌのような4足歩行動物に類似された形状・構造のロボット装置、すなわち「ペット・ロボット」においては、ユーザからの「伏せ」という命令(例えば音声入力)に応答して伏せの姿勢をとったり、自分の口の前にユーザが手を差し出すと「お手」の動作をしたりする。

【0005】このようにイヌやヒトなどの現実の動物を模したロボット装置等は、できるだけ本物の動物がする動作や感情表現に近いことが望ましい。また、ユーザからの指令や外部環境に基づいてあらかじめ決められた動作を行うだけでなく、本物の動物と同じように自律的に動作することが望ましい。何故ならば、ロボット装置等が現実とはかけ離れた同じ動作を繰り返しているとユーザは飽きてしまい、また、同じ住環境下での人間との共生という、ロボット装置等の開発の究極目的を達成し得ないからである。

【0006】最近のインテリジェントなロボット装置等は、音声入出力、音声認識、音声合成などの機能を搭載し、音声ベースでユーザと会話・対話をを行うことができる。この会話・発生においても、ユーザからの指令や外部環境に基づいて定められた動作を行うだけでなく、本物の動物と同じように自律的に動作することが望まれている。

【0007】ところで、従来の音声認識機能を有するロボット装置等は、あらかじめ設定された特定の指令語が発せられたときに、あらかじめプログラムされた特定の

動作を実行するものがほとんどであった。

【0008】この場合、ロボット装置等と対話できる語や応答文が限定されているため、言い換えれば、ロボット装置等と対話したり指令できる動作に発展性がない。

【0009】他方、インテリジェントなロボット装置等は、感情などからなる内部状態を備えており、この内部状態を外部に出力することによって、ユーザとより深いレベルでのコミュニケーションを実現することができるということが知られている。

【0010】ロボット装置等の内部状態には、「成長」や「学習」などといった、時間の経過や経験の度合いに応じて変容する要素が含まれている。ロボット装置等が、成長レベルや学習レベルなどの内部状態に応じて、外界に発現する動作パターンを変化させていくことにより、ユーザにとって飽きのこない、あるいは深い愛情を覚えることができる。言い換えれば、ロボット装置等のエンターテインメント性が向上する。

【0011】しかしながら、ユーザとの対話において成長の度合いを反映させたロボット装置等はまだ知られていない。

#### 【0012】

【発明が解決しようとする課題】本発明の目的は、ユーザとの対話機能を持つ、優れた脚式移動ロボット及びその制御方法を提供することにある。

【0013】本発明の更なる目的は、内部状態の変化に応じた対話をを行うことができる、優れた脚式移動ロボット及びその制御方法を提供することにある。

【0014】本発明の更なる目的は、内部状態の成長の度合いに応じて使用する語彙や応答のパターンなどを変化させて対話機能を成長させることができる、優れた脚式移動ロボット及びその制御方法を提供することにある。

#### 【0015】

【課題を解決するための手段及び作用】本発明は、上記課題を参照してなされたものであり、その第1の側面は、駆動可能な複数の関節角からなる脚式移動ロボット又はその制御方法であって、ユーザの音声を認識する音声認識手段又はステップと、音声認識された発話文に対する前記脚式移動ロボットの知性レベルに相応した応答文を生成する会話処理手段又はステップと、生成された応答文を音声出力する音声合成手段又はステップと、を具備することを特徴とする脚式移動ロボット又はその制御方法である。

【0016】前記会話処理手段又はステップは、知性レベルに対応付けられた単語辞書、知性レベルに対応付けられた構文解析用文法、並びに知性レベルに対応付けられた会話生成規則に基づいて、知性レベルに相応した応答文を生成する。

【0017】したがって、本発明の第1の側面に係る脚式移動ロボット又はその制御方法によれば、脚式移動ロ

ボットの内部状態の成長に合わせてユーザとの対話内容を成長させていくことができる。

【0018】例えば、脚式移動ロボットが応答文として使用する語彙や応答のパターンを成長の度合いに合わせて変化させることができる。また、事前に設定した応答発生の確率を成長に応じて変化させていくことで、最初（すなわち成長の初期）には簡単な文章しか話さないし理解できないが、成長又は時間の経過とともに徐々に複雑な発話文に対して複雑な応答文を返すことができるようになる、という印象をユーザに与えることができる。この結果、ユーザには飽きが来ない、エンターテインメント性の高い脚式移動ロボットを提供することができる。

【0019】さらに、このような成長型の対話機能を学習機能と組み合わせることによって、最初は身体（機体）の部分の名前しか覚えられることができないが、見ているオブジェクトの名前など身体とは関係ない対象物を徐々に覚えることができるよう構成することができる。

【0020】前記会話処理手段又はステップは、認識された発話文を構成する単語及び／又は構文に対応付けられた知性レベルを基に、該発話文の知性レベルを求めるようにしてもよい。

【0021】また、前記会話処理手段又はステップは、認識された発話文の知性レベルと前記脚式移動ロボットの成長度の相対値に基づいて応答文を生成するようにしてもよい。例えば、認識された発話文の知性レベルと前記脚式移動ロボットの成長度の相対値に近い知性レベルの応答文を生成するようにしてもよい。

【0022】このような場合、ユーザからの発話文の知性レベルよりも脚式移動ロボットの成長度が低い場合には、「分らない」のような応答を必ず選択せざるなど、脚式移動ロボットの成長度に応じて知性レベルの高い文章を理解できるようにするという効果を与えることができる。

【0023】また、前記会話処理手段又はステップは、認識された発話文の知性レベルと前記脚式移動ロボットの成長度の相対値に応じた各応答文の発生確率に従って応答文を選択するようにしてもよい。

【0024】このような場合、脚式移動ロボットの成長度に応じて複雑な応答が発生する確率が高くなるように応答文を生成することができる。例えば、知性レベルの値が近い応答をたくさん用意しておけば、ある成長度のときに同じ会話文を入力しても確率に応じたさまざまな応答を返すことができ、より自然な応答すなわちユーザとの対話をを行うことができる。

【0025】また、成長度に応じて応答が一意に決まらず、脚式移動ロボットの応答文に意外性を持たせることができるので、ユーザにとっては飽きの来ない、エンターテインメント性の高い脚式移動ロボットを提供するこ

とができる。

【0026】また、前記会話処理手段又はステップは、応答文の知性レベルと前記脚式移動ロボットの成長度に基づいて応答文を選択するようにしてもよい。例えば、前記脚式移動ロボットの成長度に最も近い知性レベルを持つ応答文を選択するようにしてもよい。

【0027】このような場合、脚式移動ロボットの成長度と応答文の知性レベルを直接比較するので、入力文の知性レベルを計算する必要がなく、演算負荷を軽減することができる。

【0028】また、前記会話処理手段又はステップは、前記脚式移動ロボットの成長度に応じた各応答文の発生確率に従って応答文を選択するようにしてもよい。

【0029】このような場合、脚式移動ロボットの成長度が一定の知性レベルに到達するまでは、入力文が理解できないといった幼稚な内容の応答文をより高い確率で発生させる。また、成長するにつれて、より高い知性レベルの応答文の発生確率を次第に上げていき、よい兆度に応じた知性のある対話を期待させることができる。

【0030】また、本発明の第2の側面は、ユーザからの発話文に対してキャラクタの成長度に応じた応答文を生成する対話制御方法であって、ユーザの音声を認識する音声認識ステップと、音声認識された発話文に対する前記キャラクタの知性レベルに相応した応答文を生成する会話処理ステップと、生成された応答文を音声出力する音声合成ステップと、を具備することを特徴とする対話制御方法である。

【0031】本発明の第2の側面に係る対話制御方法は、脚式移動ロボット以外に、コンピュータ・グラフィックス（CG）で動作するアニメーション・キャラクタなど、ユーザとコンピュータとの対話処理を行うさまざまな自動化装置に対して適用することができる。そして、前記会話処理手段ステップでは、知性レベルに対応付けられた単語辞書、知性レベルに対応付けられた構文解析用文法、並びに知性レベルに対応付けられた会話生成規則に基づいて、知性レベルに相応した応答文を生成するようとする。

【0032】したがって、本発明の第2の側面に係る対話制御方法によれば、キャラクタの内部状態の成長に合わせてユーザとの対話内容を成長させていくことができる。

【0033】例えば、キャラクタが応答文として使用する語彙や応答のパターンを成長の度合いに合わせて変化させることができる。また、事前に設定した応答発生の確率を成長に応じて変化させていくことで、最初（すなわち成長の初期）には簡単な文章しか話さないし理解できないが、成長又は時間の経過とともに徐々に複雑な発話文に対して複雑な応答文を返すことができるようになる、という印象をユーザに与えることができる。この結果、ユーザには飽きが来ない、エンターテインメント性

の高いユーザ・インターフェースを提供することができ

る。

【0034】また、本発明の第3の側面は、駆動可能な複数の関節角からなる脚式移動ロボットがユーザと対話するための処理をコンピュータ・システム上で実行するように記述されたコンピュータ・ソフトウェアをコンピュータ可読形式で物理的に格納した記憶媒体であって、前記コンピュータ・ソフトウェアは、ユーザの音声を認識する音声認識ステップと、音声認識された発話文に対する前記脚式移動ロボットの知性レベルに相応した応答文を生成する会話処理ステップと、生成された応答文を音声出力する音声合成ステップと、を具備することを特徴とする記憶媒体である。

【0035】また、本発明の第4の側面は、ユーザからの発話文に対してキャラクタの成長度に応じた応答文を生成する対話処理をコンピュータ・システム上で実行するように記述されたコンピュータ・ソフトウェアをコンピュータ可読形式で物理的に格納した記憶媒体であって、前記コンピュータ・ソフトウェアは、ユーザの音声を認識する音声認識ステップと、音声認識された発話文に対する前記キャラクタの知性レベルに相応した応答文を生成する会話処理ステップと、生成された応答文を音声出力する音声合成ステップと、を具備することを特徴とする記憶媒体である。

【0036】本発明の第3又は第4の各側面に係る記憶媒体は、例えば、様々なプログラム・コードを実行可能な汎用コンピュータ・システムに対して、コンピュータ・ソフトウェアをコンピュータ可読な形式で提供する媒体である。このような媒体は、例えば、CD（Compact Disc）やFD（Floppy Disk）、MO（Magneto-Optical disc）などの着脱自在で可搬性の記憶媒体である。あるいは、ネットワーク（ネットワークは無線、有線の区別を問わない）などの伝送媒体などを経由してコンピュータ・ソフトウェアを特定のコンピュータ・システムに提供することも技術的に可能である。

【0037】このような記憶媒体は、コンピュータ・システム上で所定のコンピュータ・ソフトウェアの機能を実現するための、コンピュータ・ソフトウェアと記憶媒体との構造上又は機能上の協働的関係を定義したものである。換言すれば、本発明の第3又は第4の各側面に係る記憶媒体を介して所定のコンピュータ・ソフトウェアをコンピュータ・システムにインストールすることによって、コンピュータ・システム上では協働的作用が發揮され、本発明の第1又は第3の各側面に係る脚式移動ロボット又はその制御方法、あるいは対話制御方法と同様の作用効果を得ることができる。

【0038】本発明のさらに他の目的、特徴や利点は、後述する本発明の実施例や添付する図面に基づくより詳細な説明によって明らかになるであろう。

【0039】

**【発明の実施の形態】**本発明に係る脚式移動ロボットは、内部状態の成長に応じて成長する対話機能を備えている。このような対話機能は、脚式移動型のロボット装置以外にも、その他のタイプの多関節ロボットや、コンピュータ・グラフィックス(CG)で動作するキャラクタを用いたアニメーションなど、音声による認識機能を持ち、あらかじめプログラミングされた基本動作群の中から特定の動作を音声ベースで指示することができるあらゆる動作システムに対して適用することができる。但し、以下では、自律的に基本動作を選択することができる動物型(又はペット型)ロボットを例にとって、図面を参照しながら本発明の実施例を詳解する。

【0040】図1には、本発明を実施に供される、四肢による脚式歩行を行う移動ロボット1の外観構成を示している。図示の通り、この移動ロボット1は、四肢を有する動物の形状や構造をモデルにして構成された多関節型の移動ロボットである。とりわけ本実施例の移動ロボット1は、愛玩動物の代表例であるイヌの形状及び構造を模してデザインされたペット型ロボットという側面を有し、例えば人間の住環境において人間と共に存するとともに、ユーザ操作に応答した動作表現することができる。

【0041】移動ロボット1は、胴体部ユニット2と、頭部ユニット3と、尻尾4と、四肢すなわち脚部ユニット6A～6Dで構成される。

【0042】胴体部ユニット2には、機体動作を統括的にコントロールする制御ユニット(後述:図1には図示せず)や、機体の主電源であるバッテリ21(図1には図示しない)が収容されている。

【0043】頭部ユニット3は、ロール、ピッチ及びヨーの各軸方向(図示)の自由度を持つ首関節7を介して、胴体部ユニット2の略前上端に配設されている。また、頭部ユニット3には、イヌの「目」に相当するCCD(Charge Coupled Device:電荷結合素子)カメラなどの画像認識部15と、「耳」に相当するマイクロフォン16と、「口」に相当するスピーカ17と、触感に相当するタッチ・センサ18と、複数のLEDインジケータ19と、ユーザからリモート・コントローラ(図示しない)を介して送信される指令を受信するリモート・コントローラ受信部20が搭載されている。これら以外にも、生体の五感を構成するセンサを含んでいても構わない。

【0044】尻尾4は、ロール及びピッチ軸の自由度を持つ尻尾関節8を介して、胴体部ユニット2の略後上端に湾曲若しくは揺動自在に取り付けられている。

【0045】脚部ユニット6A及び6Bは前足を構成し、脚部ユニット6C及び6Dは後足を構成する。各脚部ユニット6A～6Dは、それぞれ、大腿部ユニット9A～9D及び脛部ユニット10A～10Dの組み合わせで構成され、胴体部ユニット2の底面では、前後左右の

各隅部に取り付けられている。大腿部ユニット9A～9Dは、ロール、ピッチ、ヨーの各軸の自由度を持つ股関節11A～11Dによって、胴体部ユニット2の各々の所定部位に連結されている。また、大腿部ユニット9A～9Dと脛部ユニット10A～10Dの間は、ロール及びピッチ軸の自由度を持つ膝関節12A～12Dによって連結されている。

【0046】図示のように構成された脚式移動ロボット1は、後述する制御部からの指令により各関節アクチュエータを駆動することによって、例えば、頭部ユニット3を上下左右に振らせたり、尻尾4を振らせたり、各足部ユニット6A～6Dを同期協調的に駆動させて歩行や走行などの動作を実現することができる。

【0047】なお、移動ロボット1の関節自由度は、実際には各軸毎に配備され関節アクチュエータ(図1には図示しない)の回転駆動によって提供される。また、移動ロボット1が持つ関節自由度の個数は任意であり、本発明の要旨を限定するものではない。

【0048】図2には、この移動ロボット1の動作を制御する制御ユニットのハードウェア構成を模式的に示している。

【0049】頭部3は、マイク16及びリモート・コントローラ受信部20からなるコマンド受信部30と、画像認識部15及びタッチ・センサ18からなる外部センサ31と、スピーカ17と、LEDインジケータ19とを有している。

【0050】また、胴体部2は、主電源としてのバッテリ21を収容するとともに、バッテリ21の残存容量を検出するためのバッテリ・センサ33と、機体内部で発生する熱を検出するための熱センサ34からなる内部センサ35と、移動ロボット1全体の動作を統括的にコントロールする制御システム100とを備えている。

【0051】さらに、移動ロボット1の各部(例えば駆動関節部分)には、複数のアクチュエータ23A～23Nがそれぞれ配設されている。

【0052】コマンド受信部30は、リモート・コントローラ受信部20やマイク16などで構成され、ユーザが移動ロボット1に与える発話文を受信することができる。

【0053】リモート・コントローラ受信部20は、ユーザ操作によりリモート・コントローラ(図示しない)に入力された指令を受信して、受信信号S1Aを生成して、これを制御システム100に送出する。例えば、リモート・コントローラからの指令は、赤外線データ通信(IrDA)や、blue toothなどの近距離無線データ通信などの技術を用いて転送される。

【0054】マイク16は、ユーザの発話文を集音して音声信号S1Bを生成して、制御システム100に送出する。

【0055】外部センサ31のタッチ・センサ18は、

ユーザから移動ロボット1への働きかけ、例えば「なでる」、「たたく」などの働きかけを検出する。例えば、ユーザによりタッチ・センサ18が触れられて所望の働きかけがなされると、これに応じた接触検出信号S2Aを生成して、制御システム100に送出する。

【0056】外部センサ31の画像認識部15は、移動ロボット1の周囲の環境を識別した結果、例えば「暗い」、「お気に入りのオモチャがある」などの周囲の環境情報を検出したり、あるいは、「他のロボットが走っている」などの他の移動ロボットの動きを検出する。この画像認識部15は、周囲の画像を撮影した結果得られる画像信号S2Bを制御システム100に送出する。

【0057】内部センサ35は、移動ロボット1自身の内部状態、例えば、バッテリ容量が低下したことを意味する「お腹が空いた」、「熱がある」などの内部状態を検出するためのものであり、本実施形態ではバッテリ・センサ33と熱センサ34を含んでいる。

【0058】バッテリ・センサ33は、機体の主電源であるバッテリ21の残存容量を検出して、その結果であるバッテリ容量検出信号S3Aを制御システム100に送出する。また、熱センサ34は、機体内部の熱を検出して、その結果である熱検出信号S3Bを制御システム100に送出する。

【0059】制御システム100は、コマンド受信部30から供給される指令信号S1と、外部センサ31から供給される外部情報信号S2と、内部センサ35から供給される内部情報信号S3に基づいて、各アクチュエータ23A～23Nを駆動させるための制御信号S5A～S5Nを生成して、これらをそれぞれのアクチュエータ23A～23Nに向けて送出することによって、機体の動作を統括的にコントロールする。

【0060】本実施形態では、制御システム100は、内部状態の成長に合わせて、対話内容（あるいはその他の機体上の動作）を成長させることができる。この結果、ユーザの発話文に対する応答文などにおいて、移動ロボット1の成長性を表現することができる。また、移動ロボット1における応答文の内容が変化するので、ユーザとの対話に新鮮味を与えることができ、ユーザに飽きを感じさせない対話を実現することができる。

【0061】また、制御システム100は、外部に出力するための音声信号S10や発光信号S11を必要に応じて生成する。このうち音声信号S10をスピーカ17を介して外部に出力し、また、発光信号S11をLEDインジケータ19に送出して、所望の発光出力（例えば、点滅したり色を変化させたりする）を行うことにより、ユーザに機体の内部状態などの情報をフィードバックすることができる。例えば、発光出力により、ユーザに自己の感情を知らせるようにすることができる。なお、LEDインジケータ19に替えて、画像を表示する画像表示部（図示しない）を備えるようにしてもよい。

これにより、所望の画像表示を介して、ユーザに自己の感情などの情報をより正確且つ詳細に提示することができる。

【0062】次いで、制御システム100による移動ロボット1の音声ベースの対話メカニズムについて説明する。

【0063】図3には、制御システム100によるプログラム実行によって実現される音声ベースの対話メカニズムの機能構成を模式的に示している。同図に示すように、対話メカニズムは、音声認識部101と、会話処理部102と、音声合成部103とで構成される。

【0064】音声認識部101は、マイク16などの音声入力装置から入力された音声（ユーザの発話文など）に対して、認識結果として発話内容のテキストとその他の付随する情報を出力する。

【0065】会話処理部102は、音声認識部101が出力する情報を解析して、移動ロボット1の成長度合いを基に適切な応答文を生成する。

【0066】音声合成部103は、主として機械会話結果として得られるテキストに対して音声合成を行い、合成音声を出力する。

【0067】図4には、音声認識部101の機能構成を模式的に示している。

【0068】ユーザより発話された音声は、マイク16を介して取り込まれ、AD変換部101-1によってサンプリングされデジタル・データに変換される。

【0069】このデジタル・データは、特徴抽出部101-2において、適当なフレーム毎のスペクトル・データなどの音声の音響的な特徴を表すパラメータに変換される。

【0070】そして、比較照合部101-3では、音響モデル101-4、辞書101-5、並びに文法101-6などのデータベース情報を基に、入力されたパラメータとの比較照合処理を行って、音声認識結果として出力する。

【0071】ここで、音響モデル101-4は、対象となる言語（例えば日本語）において用いられる個々の音素や音節などの音響的な特徴を保持するモデルであり、隠れマルコフ・モデル（HMM：Hidden Markov Model）などが用いられる。（隠れマルコフ・モデルは、音素や単語の内部での特徴の時間的な変化を幾つかの状態で表現して、その状態間の遷移と、各状態でのスペクトル特徴量の様子を確率的にモデル化したものである。隠れマルコフ・モデルは、音声信号の変動を確率的に取り扱うので、DP（Dynamic Programming）マッチング法に比べて、入力音声の揺らぎなどを好適に表現することができる。）

【0072】また、辞書101-5は、認識対象となる個々の単語の発音に関する情報を記述したデータベースである。

【0073】また、文法101-6は、辞書に記述されている個々の単語がどのように連鎖し得るかを記述したデータベースであり、文脈自由文法（CFG）に基づく記述や、統計的な単語連鎖確率（N-gram）などが用いられる。

【0074】図5には、会話処理部102の機能構成を模式的に示している。

【0075】音声認識部101により音声認識結果が得られると、言語処理部102-1は、辞書102-3と解析用文法102-4を用いて、形態素解析、構文解析などの解析を認識結果に対して施し、音声認識結果を構成する単語の情報や構文情報などの言語情報を抽出する。

【0076】また、言語処理部102-1は、辞書に記述された内容を基に、ユーザから入力された音声発話の意味、意図なども抽出する。音声認識結果は、自然言語のテキストとして与えられる場合もあるし、複数の認識候補を含む文字列、あるいはグラフ構造で与えられる場合もある。

【0077】会話生成部102-2は、発話から抽出された意味情報などに基づき、会話生成規則102-5を参照して適切な応答文を生成する。

【0078】ここで、辞書102-3は、各単語の表記や解析用文法を適用するために必要な品詞情報などの情報、単語の個別の意味情報を保持している。

【0079】また、解析用文法102-4は、辞書102-3に記述された各単語の情報を基に、単語連鎖に関する制約を記述したもので、それを基に入力されたテキストを解析することができる。具体的には、正規文法や文脈自由文法、統計的な単語連鎖確率などが用いられる。また、意味的な解析までを含む場合には、HPSGなどの意味論を含んだ言語理論に基づく文法が用いられる。

【0080】また、会話生成規則102-5は、どのような応答文をどのように生成するかというルールを記述したデータベースである。

【0081】図6には、音声合成部103の機能構成を模式的に示している。

【0082】音声合成したいテキスト（すなわち、会話処理部102により生成された頭部）が与えられると、まず、テキスト解析部103-1が、辞書103-4と解析用文法103-5を用いて、形態素解析、構文解析などの解析を入力テキストに対して施し、後段の規則合成で必要となる情報を抽出する。

【0083】ここで、規則合成で必要となる情報としては、ポーズの位置に関する情報、アクセントやイントネーションを制御する情報、各単語の発音に関する情報が挙げられる。

【0084】規則合成部103-2は、これらの情報に基づいて、音素片データを接続し、ポーズ、アクセント

ト、イントネーションなどの韻律情報を適切に付与して、該当する音声のデジタル・データを生成する。

【0085】そして、DA変換部103-3は、デジタル形式の音声データをアナログ・データに変換して、音声出力を行う。

【0086】辞書103-4は、各単語の品詞情報や読み、アクセントに関する情報を保持するデータベースである。

【0087】また、解析用文法103-5は、辞書103-4に記述された各単語の情報を基に、単語連鎖に関する制約を記述したもので、それを基に入力テキストを解析することができる。

【0088】また、音素片データ103-6は、規則的に接続する音の素片のデータを保持したデータベースである。音の素片としては、例えばC V連鎖（子音母音連鎖）やV C V連鎖、C V C連鎖などの形式で保持する。

【0089】次いで、本発明に係る音声ベースの対話メカニズムについて説明する。

【0090】以下では、移動ロボット1の時刻tにおける成長レベルを $\varphi(t)$  ( $0 \leq \varphi(t) \leq 100$ ) とし、辞書102-3に登録されている単語の集合をW ( $w \in W$ ) とし、解析用文法102-4に登録されている構文の集合をC ( $c \in C$ ) とする。ここで、W, Cの各要素に知性レベルI(w)とI(c)を対応付けておく。I(w)は単語wの知性レベルを示し、また、I(c)は構文cの知性レベルを示す。

【0091】知性レベルは、単語及び構文を移動ロボット1が理解するのに必要な知性を数値化して表したものである。例えば、以下の基準に基づいて0~100の範囲で知性レベルを設定することができる。

【0092】(1) 単語の示す意味内容が単純なものほど知性レベルが低く、複雑なものほど知性レベルが高い。

(2) 単語の意味する対象物が移動ロボット1に容易に理解できるほど、知性レベルは低い。(ex. 移動ロボット1の手足などは知性レベルが低い。)

(3) 高度な視覚処理による認識能力を要するものは、知性レベルが高い。

(4) 構文については、その複雑さに応じて知性レベルを変化させる。

【0093】ある会話文sが入力され、言語処理部102-1において、語 $w_1, w_2, \dots, w_{nw}$ が単語として認識されるとともに、構文 $c_1, c_2, \dots, c_{nc}$ が文法解析結果として認識されたとき、その入力会話文sの知性レベルI(s)は、例えば次式によって求めることができる。

【0094】

【数1】

$$I(s) = \frac{1}{n_w} \sum_{i=1}^{n_w} k_w I(w_i) + \frac{1}{n_c} \sum_{j=1}^{n_c} k_c I(c_j)$$

【0095】但し、上式において、 $k_w$ 並びに $k_c$ はそれぞれ語と構文の重み係数である。

【0096】会話生成部102-2では、発話から抽出された意味情報と移動ロボット1の成長度 $\varphi(t)$ と会話文sの知性レベル $I(s)$ を用いて、応答文を生成する。

【0097】認識された会話文sに対する応答 $r_i(s)$ ( $i = 0, \dots, nr$ )が1以上登録されているとき、その知性レベル $I(r_i(s))$ をその応答内容の複雑さに応じて設定する。この設定は、応答文自体に対応付けてもよいし、上記の式【数1】を用いて応答文を構成する単語と構文から計算してもよい。値の意味については、以下に説明する2種類の方法によって異なるものとなる。また、 $r_0(s)$ は入力文が理解できなかつたことを示す「分らない」といった応答内容とし、知性レベルを最低値 $I(r_0(s)) = 0$ とする。

【0098】以下では、知性に関する3つのパラメータ、すなわち、移動ロボット1の成長度 $\varphi(t)$ 、文章sの知性レベル $I(s)$ 、応答文の知性レベル $I(r_i(s))$ を用いて移動ロボット1の成長に応じた応答を選択する方法を2種類説明する。

【0099】【方法1】まず、文章sの知性レベル $I(s)$ に対する移動ロボット1の成長度 $\varphi(t)$ の相対値(超越度) $H(t, s) = \varphi(t) - I(s)$ に応じて選択する方法について説明する。

【0100】ここでは、入力文sの知性レベル $I(s)$ に対する移動ロボット1の成長度 $\varphi(t)$ の相対値を超越度 $H(t, s)$ と定義し、 $H(t, s)$ と $I(r_i(s))$ の距離が最も近いものを選択する。すなわち、以下の式を満たす応答 $r_i(s)$ を選択する。

【0101】

【数2】

$$\min |I(r_i(s)) - H(t, s)|$$

【0102】この方法によれば、 $H(t, s) \leq 0$ のとき、すなわち入力文sの知性レベル $I(s)$ より移動ロボット1の成長度 $\varphi(t)$ が低い場合は、 $\min |I(r_i(s)) - I(r_0(s))| = I(r_0(s))$ である「分らない」という応答が必ず選択されるので、移動ロボット1が自分の成長度 $\varphi(t)$ に応じて知性レベルの高い文章を理解するという効果を持たせることができる。応答文の知性レベル $I(r_i(s))$ の設定値の範囲は、 $0 \leq I(r_i(s)) \leq 100 - I(s)$ となり、入力文sの知性レベル $I(s)$ に依存する。

【0103】上述した方法では、移動ロボット1の成長度 $\varphi(t)$ の値に応じて応答が一意に定まってしまい、応答が单调になりかねない。そこで、 $\min |I(r_i(s)) - H(t, s)|$ の代わりに以下の式を用いて、成長度 $\varphi(t)$ に対する応答 $r_i(s)$ の発生確率密度 $p(r_i(s))$ を求ることにより、応答 $r$

$i(s)$ の発生確率 $P(r_i(s))$ を求めることができる。但し、 $\sigma$ はガウス分布の大きさを表す分散値である。

【0104】

【数3】

$$p(r_i(s)) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{|I(r_i(s)) - H(t, s)|^2}{2\sigma^2}\right)$$

【0105】

【数4】

$$P(r_i(s)) = \frac{p(r_i(s))}{\sum_{i=0}^{nr} p(r_i(s))}$$

【0106】各応答 $r_i$ の発生確率 $P(r_i(s))$ と乱数発生器を組み合わせれば、移動ロボット1の成長度 $\varphi(t)$ に応じて複雑な応答が発生する確率が高くなるような応答を生成することができる。この場合、知性レベルの値が近い応答を数多く用意しておけば、同じ成長度の移動ロボット1に同じ会話文を入力しても確率に応じたさまざまな応答が返ってくることになり、より自然な応答を生成することができる。

【0107】【方法2】次いで、文章sの知性レベル $I(s)$ には無関係に、応答文 $r_i(s)$ の知性レベル $I(r_i(s))$ と移動ロボット1の成長度 $\varphi(t)$ のみで応答文を選択する方法について説明する。

【0108】ここでは、移動ロボット1の成長度 $\varphi(t)$ に最も近い知性レベル $I(r_i(s))$ を持つ応答が、以下に式に従い選択される。

【0109】

【数5】

$$\min |I(r_i(s)) - \varphi(t)|$$

【0110】この方法では、移動ロボット1の成長度 $\varphi(t)$ と応答文の知性レベルを直接比較することになるので、入力文sの知性レベル $I(s)$ を計算する必要がない。応答文 $r_i(s)$ の知性レベル $I(r_i(s))$ の設定値の範囲は、 $0 \leq I(r_i(s)) \leq 100$ である。

【0111】この方法においても、以下の式を用いれば、応答文 $r_i(s)$ の発生確率を求めることができます。

【0112】

【数6】

$$p(r_i(s)) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{|I(r_i(s)) - \varphi(t)|^2}{2\sigma^2}\right)$$

【0113】

【数7】

$$P(r_i(s)) = \frac{p(r_i(s))}{\sum_{i=0}^{n_r} p(r_i(s))}$$

【0114】最後に、上述した【方法1】において確率密度を用いる方法を適用した具体例について説明しておく。

【0115】「トマトは好き？」という会話文sが音声入力され、各パラメータや知性レベルが以下のように設定されていたとする。

【0116】

【数8】

$$\begin{aligned} r_0 &= "tomato\_ha\_siranaiyo", I(r_0(s)) = 0 \\ r_1 &= "tomato\_ha\_sukidayo", I(r_1(s)) = 20 \\ r_2 &= "tomato\_ha\_sukidakedo\_tomato\_juice\_ha\_kirai", I(r_2(s)) = 50 \end{aligned}$$

【0119】このような設定において、移動ロボット1の成長度を $0 \leq \varphi(t) \leq 100$ の範囲内で変化させて計算したそれぞれの応答文 $r_0(s)$ ,  $r_1(s)$ ,  $r_2(s)$ の確率密度 $p(r_0(s))$ ,  $p(r_1(s))$ ,  $p(r_2(s))$ をプロットしたグラフを図7に示しておく。また、それぞれの応答文 $r_0(s)$ ,  $r_1(s)$ ,  $r_2(s)$ の確率 $P(r_0(s))$ ,  $P(r_1(s))$ ,  $P(r_2(s))$ をプロットしたグラフを図8に示しておく。

【0120】図8を参照して判るように、移動ロボット1の成長度 $\varphi(t)$ が入力文sの知性レベル $I(s)$ に近づくまでは、入力文sが理解できないという内容の応答 $r_0(s)$ ＝「トマトは知らないよ。」が高い確率で発生する。

【0121】そして、時刻tの経過とともに移動ロボット1が成長するにつれて、応答文 $r_0(s)$ の発生確率が下がり、次の成長段階の応答文 $r_1(s)$ ＝「トマトは好きだよ。」の発生確率が上がる。さらに成長すると、応答文 $r_1(s)$ の発生確率が下がるにつれて、次の成長段階の応答文 $r_2(s)$ ＝「トマトは好きだけど、トマト・ジュースは嫌い。」の発生確率が高くなる。

【0122】このように確率密度に従って応答を選択することにより、移動ロボット1の成長度 $\varphi(t)$ に応じた知性のある応答文を発生する確率が高くなることが期待される。併せて、同じ成長度の移動ロボット1に同じ会話文を入力しても確率に応じたさまざまな応答が返ってくることになり、より自然な応答を生成することができる。

【0123】【追補】以上、特定の実施例を参照しながら、本発明について詳解してきた。しかしながら、本発

$$k_w = 0.6, k_c = 0.4$$

$$\sigma = 10$$

$$w_1 = tomato, I(w_1) = 50$$

$$w_2 = suki, I(w_2) = 30$$

$$c_1 = SVO, I(c_1) = 20$$

【0117】会話生成部102-2は、この入力文sに対する応答文の候補を生成する。生成された各応答文の内容と知性レベルを以下に例示する。

【0118】

【数9】

明の要旨を逸脱しない範囲で当業者が該実施例の修正や代用を成し得ることは自明である。

【0124】本明細書では、ペット・ロボットに対して本発明を適用した実施形態を主として説明してきたが、本発明の要旨は、必ずしも「ロボット」と称される製品には限定されない。すなわち、電気的若しくは磁気的な作用を用いて人間の動作に似せた運動を行う機械装置であるならば、例えば玩具等のような他の産業分野に属する製品であっても、同様に本発明を適用することができる。また、勿論、コンピュータ・グラフィックス(CG)で動作するキャラクタを用いたアニメーションなどに対しても本発明を適用することができる。

【0125】要するに、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本発明の要旨を判断するためには、冒頭に記載した特許請求の範囲の欄を参照すべきである。

【0126】

【発明の効果】以上詳記したように、本発明によれば、ユーザとの対話機能を持つ、優れた脚式移動ロボット及びその制御方法を提供することができる。

【0127】また、本発明によれば、内部状態の変化に応じた対話をを行うことができる、優れた脚式移動ロボット及びその制御方法を提供することができる。

【0128】また、本発明によれば、内部状態の成長の度合いに応じて使用する語彙や応答のパターンなどを変化させて対話機能を成長させることができる、優れた脚式移動ロボット及びその制御方法を提供することができる。

【0129】本発明に係る脚式移動ロボットは、内部状態の成長の度合いに合わせて、対話が成長する。例えば、使用する語彙や応答のパターンが変化する。事前に

設定した応答発生の確率を、成長に応じて変化させることで、最初は簡単な文しかしゃべらないし理解しないが、徐々に複雑な文に対して複雑な応答を行えるようになる。また、このような成長型の対話機能を学習機能と組み合わせることによって、最初は身体の部分の名前しか覚えられないが、徐々に見ているオブジェクトの名前など身体と関係ないものも覚えられるようになる。

【0130】本発明によれば、対話の内容をロボットの成長度合いに応じて変化させることにより、ロボットの成長性を表現することができる。また、対話の内容が変化することにより、ユーザとの対話に新鮮味を与えることができ、ユーザに飽きを感じさせない対話機能を実現することができる。

#### 【図面の簡単な説明】

【図1】本発明を実施に供される四肢による脚式歩行を行う移動ロボット1の外観構成を示した図である。

【図2】移動ロボット1の動作を制御する制御ユニットのハードウェア構成を模式的に示した図である。

【図3】制御システム100によるプログラム実行によって実現される音声ベースの対話メカニズムの機能構成を模式的に示した図である。

【図4】音声認識部101の機能構成を模式的に示した図である。

【図5】会話処理部102の機能構成を模式的に示した図である。

【図6】音声合成部103の機能構成を模式的に示した図である。

【図7】移動ロボット1の成長度を $0 \leq \varphi(t) \leq 10$ の範囲内で変化させて計算したそれぞれの応答文 $r_0(s), r_1(s), r_2(s)$ の確率密度 $p(r_0(s)), p(r_1(s)), p(r_2(s))$ をプロットしたグラフである。

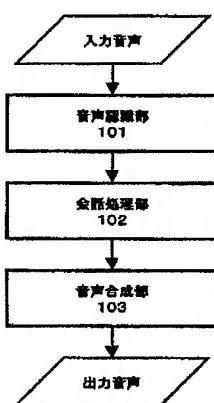
【図8】移動ロボット1の成長度を $0 \leq \varphi(t) \leq 10$

の範囲内で変化させて計算したそれぞれの応答文 $r_0(s), r_1(s), r_2(s)$ の確率 $P(r_0(s)), P(r_1(s)), P(r_2(s))$ をプロットしたグラフである。

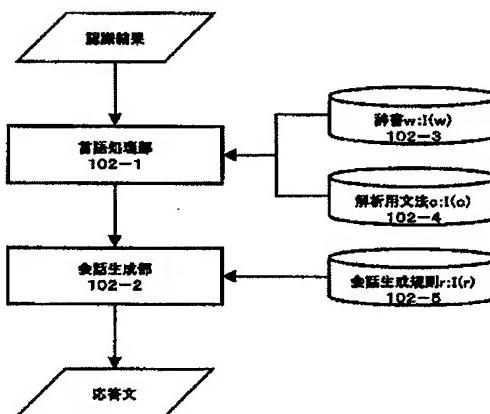
#### 【符号の説明】

- 1…移動ロボット
- 2…胴体部ユニット
- 3…頭部ユニット
- 4…尻尾
- 6 A～6 D…脚部ユニット
- 7…首関節
- 8…尻尾関節
- 9 A～9 D…大腿部ユニット
- 10 A～10 D…脛部ユニット
- 11 A～11 D…股関節
- 12 A～12 D…膝関節
- 15…カメラ（画像認識部）
- 16…マイクロフォン
- 17…スピーカ
- 18…タッチセンサ
- 19…LEDインジケータ
- 20…リモート・コントローラ受信部
- 21…バッテリ
- 23…アクチュエータ
- 30…コマンド受信部
- 31…外部センサ
- 33…バッテリ・センサ
- 34…熱センサ
- 35…内部センサ
- 100…制御システム
- 101…音声認識部
- 102…会話処理部
- 103…音声合成部

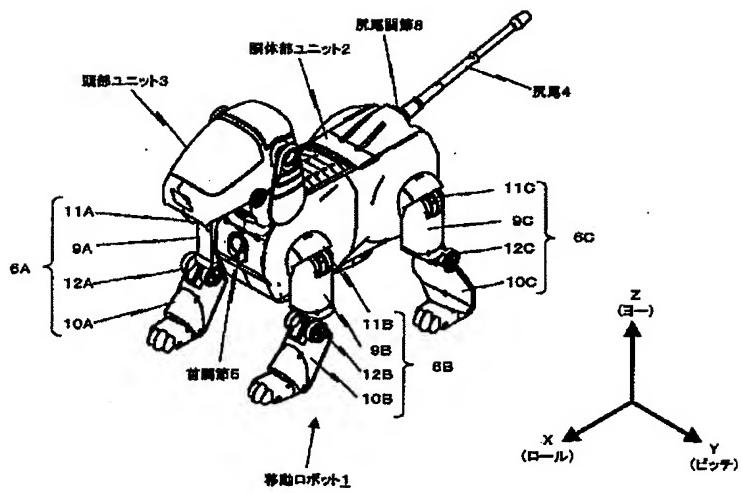
【図3】



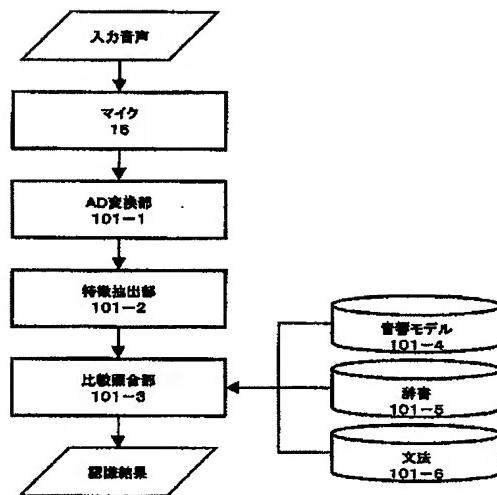
【図5】



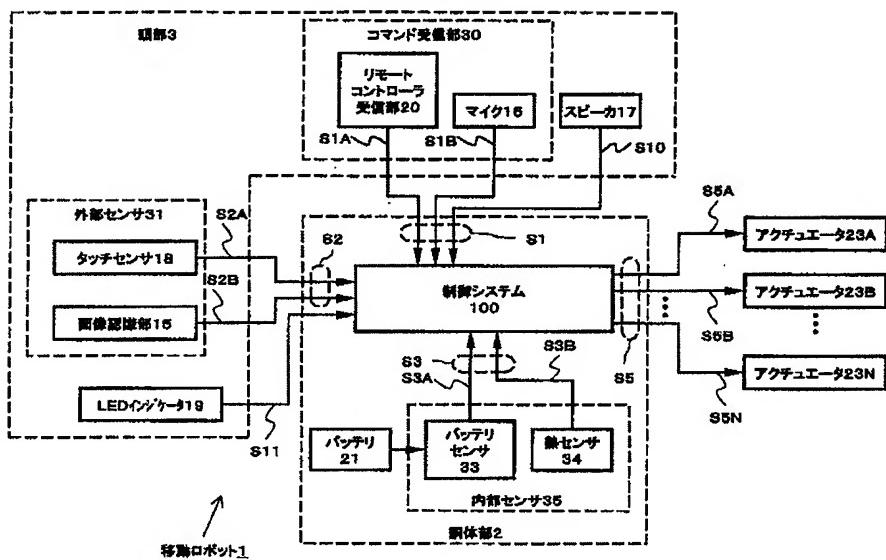
【図1】



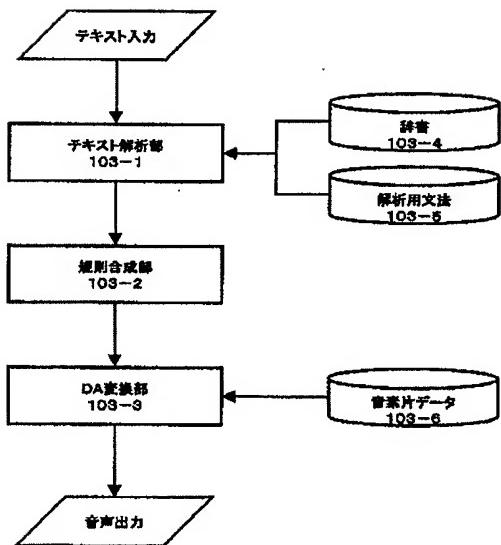
【図4】



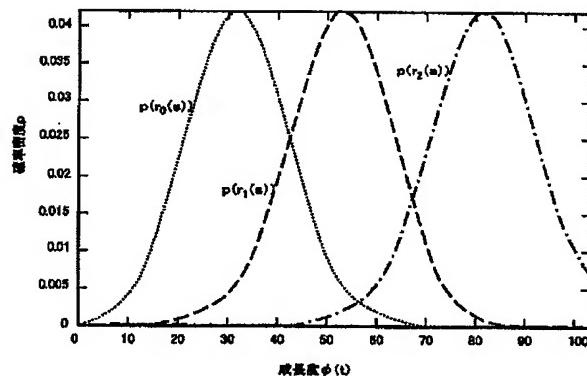
【図2】



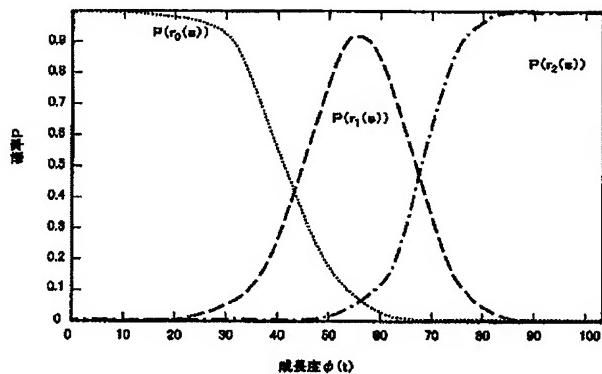
【図6】



【図7】



【図8】



## フロントページの続き

(51) Int.CI.7

G 1 O L 15/00  
15/18  
15/22

識別記号

F I  
G 1 O L 3/00

「テーマコード」(参考)

5 3 7 D  
5 3 7 G  
5 5 1 H  
R  
5 7 1 U

(72) 発明者 石井 和夫

東京都品川区北品川6丁目7番35号 ソニ  
一株式会社内

(72) 発明者 下村 秀樹

東京都品川区北品川6丁目7番35号 ソニ  
一株式会社内

(72)発明者 ウゴ ディ プロフィオ  
東京都品川区北品川6丁目7番35号 ソニ  
一株式会社内

Fターム(参考) 2C150 CA01 CA02 CA04 DA05 DA24  
DA26 DA27 DA28 DF03 DF04  
DF06 DF33 ED42 ED47 ED52  
EF03 EF07 EF13 EF16 EF23  
EF28 EF29 EF33 EF36  
3C007 AS36 CS08 KS39 WA04 WA14  
WB19 WC01 WC07  
5D015 HH21 KK02 KK04  
5D045 AB11

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-321177  
(43)Date of publication of application : 05.11.2002

---

(51)Int.Cl. B25J 5/00  
A63H 3/33  
A63H 11/00  
B25J 13/00  
G10L 13/00  
G10L 15/00  
G10L 15/18  
G10L 15/22

---

(21)Application number : 2001- (71)Applicant : SONY CORP  
124974  
(22)Date of filing : 23.04.2001 (72)Inventor : OKUBO ATSUSHI  
COSTA GABRIEL  
ISHII KAZUO  
SHIMOMURA HIDEKI  
DI PROFIO UGO

---

## (54) LEG TYPE MOVING ROBOT AND ITS CONTROL METHOD

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide a leg type moving robot that grows a communication function according to the growth level of the inside state.

**SOLUTION:** A speech recognition system has a word/syntax recognition function and associates an intelligence level to each registered word and syntax. By comparing a conversation start sentence and a response sentence with the intelligence level and adapting the growth level of the robot then the occurrence probability of a response to the recognized conversation start sentence is calculated the system changes conversation contents according to the growth level of the robot. In accordance with the growth of inside state the robot can chat and understand only easy sentences at the beginning but it comes to be possible to manipulate gradually a complicated response to a complicated sentence.

---

### CLAIMS

---

[Claim(s)]

[Claim 1]A leg formula mobile robot which consists of two or more joint angles which can be driven comprising:

A voice recognition means which recognizes a user's sound.

A conversational-processing means to generate a response sentence which \*\*\*\*ed in said leg formula mobile robot's intelligence level to an utterance sentence by which speech recognition was carried out and a voice synthesis means which carries out voice response of the generated response sentence.

[Claim 2]The leg formula mobile robot according to claim 1 characterized by what said conversational-processing means generates a response sentence which \*\*\*\*ed in an intelligence level for based on a conversation production rule matched with a word dictionary matched with an intelligence level grammar for syntax analysis matched with an intelligence level and an intelligence level.

[Claim 3]The leg formula mobile robot according to claim 1 characterized by what said conversational-processing means asks for an intelligence level of this utterance sentence based on an intelligence level matched with a word and/or syntax which constitute a recognized utterance sentence.

[Claim 4]The leg formula mobile robot according to claim 1 characterized by what said conversational-processing means generates a response sentence for based on an intelligence level of a recognized utterance sentence and a relative value of said leg formula mobile robot's growing degrees.

[Claim 5]The leg formula mobile robot according to claim 1 characterized by what said conversational-processing means generates a response sentence of an intelligence level of a recognized utterance sentence and an intelligence level near a relative value of said leg formula mobile robot's growing degrees for.

[Claim 6]The leg formula mobile robot according to claim 1 characterized by what said conversational-processing means chooses a response sentence for according to the probability of occurrence of each response sentence according to an intelligence level of a recognized utterance sentence and a relative value of said leg formula mobile robot's growing degrees.

[Claim 7]The leg formula mobile robot according to claim 1 characterized by what said conversational-processing means chooses a response sentence for based on an intelligence level of a response sentence and said leg formula mobile robot's growing degrees.

[Claim 8]The leg formula mobile robot according to claim 1 characterized by what said conversational-processing means chooses a response sentence with an intelligence level nearest to said leg formula mobile robot's growing degrees for.

[Claim 9]The leg formula mobile robot according to claim 1 characterized by what said conversational-processing means chooses a response sentence for according to the probability of occurrence of each response sentence according to said leg formula mobile robot's growing degrees.

[Claim 10]A control method of a leg formula mobile robot which consists of two or more joint angles which can be driven characterized by comprising the following.

A speech recognition step which recognizes a user's sound.

A conversational-processing step which generates a response sentence which \*\*\*\*ed in said leg formula mobile robot's intelligence level to an utterance sentence by which speech recognition was carried outand a voice synthesis step which carries out voice response of the generated response sentence.

[Claim 11]A word dictionary matched with an intelligence level in said conversational-processing stepA control method of the leg formula mobile robot according to claim 10 characterized by what a response sentence which \*\*\*\*ed in an intelligence level is generated for based on a conversation production rule matched with grammar for syntax analysis matched with an intelligence leveland an intelligence level.

[Claim 12]A control method of the leg formula mobile robot according to claim 10 characterized by what is asked for an intelligence level of this utterance sentence based on an intelligence level matched with a word and/or syntax which constitute a recognized utterance sentence from said conversational-processing step.

[Claim 13]A control method of the leg formula mobile robot according to claim 10 characterized by what a response sentence is generated for in said conversational-processing step based on an intelligence level of a recognized utterance sentenceand a relative value of said leg formula mobile robot's growing degrees.

[Claim 14]A control method of the leg formula mobile robot according to claim 10 characterized by what a response sentence of an intelligence level of a recognized utterance sentence and an intelligence level near a relative value of said leg formula mobile robot's growing degrees is generated for in said conversational-processing step.

[Claim 15]The leg formula mobile robot according to claim 10 characterized by what said conversational-processing means chooses a response sentence for according to the probability of occurrence of each response sentence according to an intelligence level of a recognized utterance sentenceand a relative value of said leg formula mobile robot's growing degrees.

[Claim 16]A control method of the leg formula mobile robot according to claim 10 characterized by what a response sentence is chosen for based on an intelligence level of a response sentenceand said leg formula mobile robot's growing degrees in said conversational-processing step.

[Claim 17]A control method of the leg formula mobile robot according to claim 10 characterized by what a response sentence which has an intelligence level nearest to said leg formula mobile robot's growing degrees in said conversational-processing step is chosen for.

[Claim 18]A control method of the leg formula mobile robot according to claim 10 characterized by what a response sentence is chosen for according to the probability of occurrence of each response sentence according to said leg formula mobile robot's growing degrees in said conversational-processing step.

[Claim 19]A dialog control method which generates a response sentence according

to growing degrees of a character to an utterance sentence from a usercomprising:

A speech recognition step which recognizes a user's sound.

A conversational-processing step which generates a response sentence which \*\*\*\*ed in an intelligence level of said character to an utterance sentence by which speech recognition was carried outand a voice synthesis step which carries out voice response of the generated response sentence.

[Claim 20]A dialog control method according to claim 19 characterized by what a response sentence which \*\*\*\*ed in an intelligence level is generated for based on a conversation production rule matched with a word dictionary matched with an intelligence levelgrammar for syntax analysis matched with an intelligence leaveland an intelligence level in said conversational-processing step.

[Claim 21]A dialog control method according to claim 19 characterized by what is asked for an intelligence level of this utterance sentence based on an intelligence level matched with a word and/or syntax which constitute a recognized utterance sentence from said conversational-processing step.

[Claim 22]A dialog control method according to claim 19 characterized by what a response sentence is generated for in said conversational-processing step based on an intelligence level of a recognized utterance sentenceand a relative value of growing degrees of said character.

[Claim 23]A dialog control method according to claim 19 characterized by what a response sentence of an intelligence level of a recognized utterance sentence and an intelligence level near a relative value of growing degrees of said character is generated for in said conversational-processing step.

[Claim 24]A dialog control method according to claim 19 characterized by what said conversational-processing means chooses a response sentence for according to the probability of occurrence of each response sentence according to an intelligence level of a recognized utterance sentenceand a relative value of growing degrees of said character.

[Claim 25]A dialog control method according to claim 19 characterized by what a response sentence is chosen for based on an intelligence level of a response sentenceand growing degrees of said character in said conversational-processing step.

[Claim 26]A dialog control method according to claim 19 characterized by what a response sentence which has an intelligence level nearest to growing degrees of said character in said conversational-processing step is chosen for.

[Claim 27]A dialog control method according to claim 19 characterized by what a response sentence is chosen for according to the probability of occurrence of each response sentence according to said leg formula mobile robot's growing degrees in said conversational-processing step.

[Claim 28]A storage which stored physically computer software described to perform processing for a leg formula mobile robot which consists of two or more joint angles which can be driven to converse with a user on computer systems in computer-

readable formcomprising:

A speech recognition step said computer software recognizes a user's sound to be.

A conversational-processing step which generates a response sentence which \*\*\*\*ed in said leg formula mobile robot's intelligence level to an utterance sentence by which speech recognition was carried outand a voice synthesis step which carries out voice response of the generated response sentence.

[Claim 29]A storage which stored physically computer software described to perform interactive processing which generates a response sentence according to growing degrees of a character to an utterance sentence from a user on computer systems in computer-readable formcomprising:

A speech recognition step said computer software recognizes a user's sound to be.

A conversational-processing step which generates a response sentence which \*\*\*\*ed in an intelligence level of said character to an utterance sentence by which speech recognition was carried outand a voice synthesis step which carries out voice response of the generated response sentence.

---

## DETAILED DESCRIPTION

---

### [Detailed Description of the Invention]

#### [0001]

[Field of the Invention]This invention relates to a leg formula mobile robot which consists of two or more joint angles which can be drivenand a method for controlling the sameand relates to a leg formula mobile robot which has interactive service with a user especiallyand a method for controlling the same.

[0002]In detailthis invention relates to a leg formula mobile robot which performs the dialog according to change of the internal stateand a method for controlling the sameand relates to a leg formula mobile robot which changes the vocabulary especially used according to the degree of growth of an internal statethe pattern of a responseetc.and grows up interactive serviceand a method for controlling the same.

#### [0003]

[Description of the Prior Art]Conventionallythe plentiful proposal is made and developed about the animation using the character which operates with the robot device of the walking type which operates according to the environment of the instructions from a useror the circumferencean articulated robotor computer graphics (CG)etc. Such a robot device or animation (belowthese are summarized and a "robot device" is called) answers the instructions from a userand performs a series of operations.

[0004]For examplein the robot device of shape and structure similar to a quadrapedalism animal like a dogi.e.a "pet robot" The posture answered and turned down on the command (for examplevoice input) "lie down" from a user is takenor

operation of a "hand" will be carried out if a user holds out a hand before his mouth.

[0005]Thusas for the robot device which imitated actual animalssuch as a dog and Homo sapiensit is desirable that it is close to the operation and the emotional expression which a real animal carries out as much as possible. It is desirable it not only to to perform operation for which it opted beforehand based on the instructions and the outside environment from a userbut to operate autonomously like a real animal. Reality is because a user gets bored and cannot attain the final end of development of a robot device etc. called symbiosis with human being under the same living conditionsif the robot device etc. have repeated the same operation that was widely different.

[0006]The latest intelligent robot device can carry functionssuch as a voice input/outputsspeech recognitionand voice synthesisand can perform conversation and a dialog with a user on a voice basis. It not only performs operation defined based on the instructions and the outside environment from a userbut to operate autonomously like a real animal is desired in this conversation and generating.

[0007]By the waywhen the specific instruction word to which robot devices which have the conventional speech recognition function were set beforehand was emittedit was almost the case which performs specific operation programmed beforehand.

[0008]In this casesince the word and response sentence which can have a dialog with a robot device etc. are limitedif it puts in another waythere will be no possibilities in the operation which can have a dialog with a robot device etc.or it can be ordered.

[0009]On the other handintelligent robot devices are provided with the internal state which consists of feeling etc.and it is known by outputting this internal state outside that communication with a user and a deeper level is realizable.

[0010]The element which changes according to the passage of time and the degrees of experiencesuch as "growth" and "study" is included in the internal state of a robot device etc. By changing the operation pattern which a robot device etc. reveal in the external world according to internal statessuch as a growth level and a learning levelthe weariness for a user does not come or deep love can be memorized. In other wordsentertainment naturesuch as a robot deviceimproves.

[0011]Howeverthe robot device in which the degree of growth was made to reflect in a dialog with a user is not known yet.

[0012]

[Problem(s) to be Solved by the Invention]The purpose of this invention is to provide a great leg formula mobile robot with interactive service with a userand a method for controlling the same.

[0013]The further purpose of this invention is to provide a great leg formula mobile robot which can perform the dialog according to change of the internal stateand a method for controlling the same.

[0014]The further purpose of this invention is to provide a great leg formula mobile

robot which can change the vocabulary used according to the degree of growth of an internal state the pattern of a response etc. and can grow up interactive service and a method for controlling the same.

[0015]

[Means for Solving the Problem and its Function] This invention is made in consideration of an aforementioned problem and the 1st side A voice recognition means or a step which is a leg formula mobile robot which consists of two or more joint angles which can be driven or its control method and recognizes a user's sound. They are a leg formula mobile robot possessing a conversational-processing means or a step which generates a response sentence which \*\*\*\*ed in said leg formula mobile robot's intelligence level to an utterance sentence by which speech recognition was carried out and a voice synthesis means or a step which carries out voice response of the generated response sentence or its control method.

[0016] Said conversational-processing means or a step generates a response sentence which \*\*\*\*ed in an intelligence level based on a conversation production rule matched with a word dictionary matched with an intelligence level grammar for syntax analysis matched with an intelligence level and an intelligence level.

[0017] Therefore according to a leg formula mobile robot concerning the 1st side of this invention or its control method dialog contents with a user can be grown up according to growth of a leg formula mobile robot's internal state.

[0018] For example a leg formula mobile robot can change a vocabulary used as a response sentence and a pattern of a response according to a degree of growth. By what probability of response generating set up a priori is changed for according to growth. Although it speaks only about an easy text first (namely early stages of growth) and he cannot understand an impression that a complicated response sentence can be gradually returned now to a complicated utterance sentence with growth or progress of time can be given to a user. As a result a user can be provided with a high leg formula mobile robot of entertainment nature by which weariness does not come.

[0019] Although only a name of a portion of the body (body) can be memorized at first by combining such grown up type interactive service with a learning function it can constitute so that a subject which is not related can be gradually learned the bodies such as a name etc. of an object currently seen.

[0020] It may be made for said conversational-processing means or a step to ask for an intelligence level of this utterance sentence based on an intelligence level matched with a word and/or syntax which constitute a recognized utterance sentence.

[0021] It may be made for said conversational-processing means or a step to generate a response sentence based on an intelligence level of a recognized utterance sentence and a relative value of said leg formula mobile robot's growing degrees. For example it may be made to generate a response sentence of an intelligence level of a recognized utterance sentence and an intelligence level near a relative value of said leg formula mobile robot's growing degrees.

[0022] In such a case when a leg formula mobile robot's growing degrees are lower

than an intelligence level of an utterance sentence from a useran effect that making a response like "not understanding" certainly choose etc. enables it to understand a text with a high intelligence level according to a leg formula mobile robot's growing degrees can be given.

[0023]It may be made for said conversational-processing means or a step to choose a response sentence according to the probability of occurrence of each response sentence according to an intelligence level of a recognized utterance sentenceand a relative value of said leg formula mobile robot's growing degrees.

[0024]In such a casea response sentence is generable so that probability that a complicated response will occur according to a leg formula mobile robot's growing degrees may become high. For exampleif a value of an intelligence level prepares much near responseeven if it inputs the conversational sentences same at the time of a certain growing degreesvarious responses according to probability can be returnedand a dialogmore natural responsei.e.usercan be performed.

[0025]Since it cannot be decided according to growing degrees that a response will be a meaning but unexpected nature can be given to a leg formula mobile robot's response sentencefor a usera high leg formula mobile robot of entertainment nature by which weariness does not come can be provided.

[0026]It may be made for said conversational-processing means or a step to choose a response sentence based on an intelligence level of a response sentenceand said leg formula mobile robot's growing degrees. For exampleit may be made to choose a response sentence with an intelligence level nearest to said leg formula mobile robot's growing degrees.

[0027]In such a casesince a leg formula mobile robot's growing degrees are directly compared with an intelligence level of a response sentenceit is not necessary to calculate an intelligence level of an input sentenceand arithmetic load can be reduced.

[0028]It may be made for said conversational-processing means or a step to choose a response sentence according to the probability of occurrence of each response sentence according to said leg formula mobile robot's growing degrees.

[0029]In such a casea response sentence of childish contents that he cannot understand an input sentence is generated in higher probability until a leg formula mobile robot's growing degrees reach a fixed intelligence level. The probability of occurrence of a response sentence of a higher intelligence level is raised graduallyand a dialog with intelligence according to 1 trillion good times can be made to expect as it grows up.

[0030]A speech recognition step which the 2nd side of this invention is the dialog control method which generates a response sentence according to growing degrees of a character to an utterance sentence from a userand recognizes a user's soundIt is the dialog control method possessing a conversational-processing step which generates a response sentence which \*\*\*\*ed in an intelligence level of said character to an utterance sentence by which speech recognition was carried outand a voice synthesis step which carries out voice response of the generated response sentence.

[0031]The dialog control method concerning the 2nd side of this invention can apply an animation character etc. which operate with computer graphics (CG) to various automated equipment which performs interactive processing of a user and a computer in addition to a leg formula mobile robot. And in said conversational-processing means stepa response sentence which \*\*\*\*ed in an intelligence level is generated based on a conversation production rule matched with a word dictionary matched with an intelligence levelgrammar for syntax analysis matched with an intelligence leveland an intelligence level.

[0032]Thereforeaccording to a dialog control method concerning the 2nd side of this inventiondialog contents with a user can be grown up according to growth of an internal state of a character.

[0033]For examplea vocabulary which a character uses as a response sentenceand a pattern of a response can be changed according to a degree of growth. By what probability of response generating set up a priori is changed for according to growth. Although it speaks only about an easy text first (namelyearly stages of growth) and he cannot understandan impression that a complicated response sentence can be gradually returned now to a complicated utterance sentence with growth or progress of time can be given to a user. As a resulta user can be provided with a high user interface of entertainment nature by which weariness does not come.

[0034]The 3rd side of this invention is the storage which stored physically computer software described to perform processing for a leg formula mobile robot which consists of two or more joint angles which can be driven to converse with a user on computer systems in computer-readable formA speech recognition step said computer software recognizes a user's sound to belt is a storage possessing a conversational-processing step which generates a response sentence which \*\*\*\*ed in said leg formula mobile robot's intelligence level to an utterance sentence by which speech recognition was carried outand a voice synthesis step which carries out voice response of the generated response sentence.

[0035]The 4th side of this invention is the storage which stored physically computer software described to perform interactive processing which generates a response sentence according to growing degrees of a character to an utterance sentence from a user on computer systems in computer-readable formA speech recognition step said computer software recognizes a user's sound to belt is a storage possessing a conversational-processing step which generates a response sentence which \*\*\*\*ed in an intelligence level of said character to an utterance sentence by which speech recognition was carried outand a voice synthesis step which carries out voice response of the generated response sentence.

[0036]A storage concerning each 3rd [ of this invention ] or 4th side is a medium which provides computer software in a computer-readable form to a general purpose computer system which can execute various program codesfor example.Attachment and detachment of CD (Compact Disc)FD (Floppy Disk)MO (Magneto–Opticaldisc)etc.etc. are free for such a mediumand it is a storage of portabilityfor example. Or it is also technically possible to provide specific

computer systems with computer software via transmission mediasuch as a network (a network does not ask distinction of radio and a cable)etc.

[0037]Such a storage defines a collaboration relation on structure of computer software and a storage for realizing a function of predetermined computer softwareor a function on computer systems. By installing predetermined computer software in computer systems via a storage concerning each 3rd [ of this invention ]or 4th sideif it puts in another wayOn computer systemsa collaboration operation is demonstrated and the same operation effect as a leg formula mobile robot concerning each 1st [ of this invention ] or 3rd sideits control methodor a dialog control method can be obtained.

[0038]The purposethe featureand an advantage of further others of this invention will become clear [ rather than ] by detailed explanation based on an example and a drawing to attach of this invention mentioned later.

[0039]

[Embodiment of the Invention]The leg formula mobile robot concerning this invention has the interactive service which grows according to growth of an internal state. Such interactive service besides a leg formula movement type robot device The articulated robot of other typesThe animation using the character which operates with computer graphics (CG)etc.It can have recognizing ability with a sound and specific operation can be applied to all the actuating systems that can be directed on a voice basis out of the basic motion group programmed beforehand. Howeverbelowthe animal type (or pet type) robot which can choose basic motion autonomously is taken for an exampleand the example of this invention is explained in detailreferring to drawings.

[0040]The appearance composition of the mobile robot 1 which performs the leg formula walk by the limbs with which operation is presented in this invention is shown in drawing 1. This mobile robot 1 is a many joint type mobile robot which used as the model the shape and structure of the animal of having the limbsand was constituted as a graphic display. while the mobile robot 1 of this example has the side of the robot pet designed by imitating the shape and structure of a dog which are the examples of representation of a petfor examplecoexisting with human being in human being's living conditionsuser's operation was especially answered -- expression of operation can be carried out.

[0041]The mobile robot 1 comprisesthe idiosoma unit 2the head unit 3the tail 4and the limbs 6A-6Di.e.leg units.

[0042]The control unit (after-mentioned: not shown to drawing 1) which controls body operation in generalizationand the battery 21 (not shown to drawing 1) which is the main power supplies of the body are accommodated in the idiosoma unit 2.

[0043]The head unit 3 is allocated in the approximately front upper bed of the idiosoma unit 2 via the head joint 7 with a rolla pitchand the flexibility of each shaft orientations (graphic display) of a yaw. The image recognition sections 15such as a CCD (Charge Coupled Device: charge coupled device) camera which is equivalent to the head unit 3 at the "eye" of a dogThe microphone 16 equivalent to an "ear"and the loudspeaker 17 equivalent to a "mouth"The <sup>remote</sup>

~~controller~~ receive section 20 which receives the touch sensor 18 equivalent to tactile feeling two or more LED indicators 19 and the instructions transmitted via a ~~remote~~ controller (not shown) from a user is carried. The sensor which constitutes a living body's senses besides these may be included.

[0044] The tail 4 is attached to the approximately back upper bed of the idiosoma unit 2 via the tail joint 8 with the flexibility of a roll and a pitch axis enabling a curve or free rocking.

[0045] The leg units 6A and 6B constitute a forefoot and the leg units 6C and 6D constitute hind legs. Each leg units 6A-6D comprise combination of the femoral region units 9A-9D and the leg part units 10A-10D and are attached to each corner of front and rear right and left on the bottom of the idiosoma unit 2 respectively. The femoral region units 9A-9D are connected with each predetermined region of the idiosoma unit 2 by the hip joints 11A-11D with a roll and the flexibility of each axis of a yaw. It is connected by the knee joints 12A-12D with the flexibility of a roll and a pitch axis between the femoral region units 9A-9D and the leg part units 10A-10D.

[0046] The leg formula mobile robot 1 constituted like a graphic display By driving each joint actuator by the instructions from a control section mentioned later the head unit 3 can be made to be able to shake vertically and horizontally or the tail 4 can be made to be able to wag or a synchronous cooperation target is made to drive each foot units 6A-6D and operation of a walka run etc. can be realized for example.

[0047] Actually the mobile robot's 1 joint flexibility is arranged for every axis and is provided by rotation of a joint actuator (not shown to drawing 1). The number of the joint flexibility which the mobile robot 1 has is arbitrary and does not limit the gist of this invention.

[0048] The hardware constitutions of the control unit which controls operation of this mobile robot 1 are typically shown in drawing 2.

[0049] The head 3 is provided with the following.

The command reception part 30 which consists of the microphone 16 and the ~~remote~~ controller receive section 20.

The external sensor 31 which consists of the image recognition section 15 and the touch sensor 18.

Loudspeaker 17.

LED indicator 19.

[0050] The idiosoma 2 is provided with the following.

The battery sensor 33 for detecting the remaining capacity of the battery 21 while accommodating the battery 21 as a main power supply.

The internal sensor 35 which consists of the heat sensor 34 for detecting the heat generated inside the body.

The control system 100 which controls operation of the mobile-robot 1 whole in generalization.

[0051]Two or more actuators 23A–23N are allocated by the mobile robot's 1 each part (for example driving joints portion) respectively.

[0052]The command reception part 30 comprises the ~~remote controller~~ receive section 20 the microphone 16 etc. and can receive the utterance sentence which a user gives to the mobile robot 1.

[0053]The ~~remote controller~~ receive section 20 receives the instructions inputted into the ~~remote controller~~ (not shown) by user's operation generates the input signal S1A and sends out \*\* to the control system 100. For example the instructions from a ~~remote controller~~ are transmitted using arts such as infrared ray data communication (IrDA) and short-distance-radio data communications such as bluetooth.

[0054]The microphone 16 collects a user's utterance sentence generates the audio signal S1B and sends it out to the control system 100.

[0055]The touch sensor 18 of the external sensor 31 detects the influence to the mobile robot 1 from a user for example influence of "it strokes" "it strikes" etc. For example if the touch sensor 18 is touched by the user and influence of a request is made the contact detecting signal S2A according to this is generated and it sends out to the control system 100.

[0056]that the image recognition section 15 of the external sensor 31 detects the environment information of the circumferences such as "it is dark" and there "there being a favorite toy" for example as a result of identifying the environment around the mobile robot 1 \*\*\*\* -- " -- others -- the robot is running -- " -- etc. -- a motion of other mobile robots is detected. This image recognition section 15 sends out picture signal S2B obtained as a result of photoing the surrounding picture to the control system 100.

[0057]The internal sensor 35 is for detecting the internal state of mobile-robot 1 self for example the internal state meaning battery capacity having fall such as "I am hungry" and "it being feverish" and the battery sensor 33 and the heat sensor 34 are included in this embodiment.

[0058]The battery sensor 33 detects the remaining capacity of the battery 21 which is a main power supply of the body and sends out the battery capacity detecting signal S3A which is the result to the control system 100. The heat sensor 34 detects the heat inside the body and sends out the heat detecting signal S3B which is the result to the control system 100.

[0059]The command signal S1 with which the control system 100 is supplied from the command reception part 30 Based on the external information signal S2 supplied from the external sensor 31 and the inside information signal S3 supplied from the internal sensor 35 control signal S5 A-S 5N for making each actuators 23A–23N drive is generated By turning these to each actuator 23A–23N and sending them out operation of the body is controlled in generalization.

[0060]In this embodiment the control system 100 can grow up dialog contents (or operation on other bodies) according to growth of an internal state. As a result the mobile robot's 1 growth possibility can be expressed in the response sentence to a user's utterance sentence etc. Since the contents of the response sentence in the mobile robot 1 change freshness can be given to a dialog with a user and the dialog

for which a user is not made to sense weariness can be realized.

[0061]The control system 100 generates the audio signal S10 and the flashing caution signal S11 for outputting outside if needed. Among theseby outputting the audio signal S10 outside via the loudspeaker 17and sending out the flashing caution signal S11 to LED indicator 19and performing a desired radiant power output (for exampleit blinks or a color is changed)Informationincluding the internal state of the bodyetc.can be fed back to a user. For examplea user can be told about self feeling by a radiant power output. It changes to LED indicator 19 and may be made to have a picture display part (not shown) which displays a picture. Therebyinformationincluding self feeling etc.can be shown to a user more correctly and in detail via desired image display.

[0062]Subsequentlythe dialog mechanism of the voice base of the mobile robot 1 by the control system 100 is explained.

[0063]The functional constitution of the dialog mechanism of a voice base realized by the program execution by the control system 100 is typically shown in drawing 3. As shown in the figurea dialog mechanism comprises the voice recognition part 101the conversational-processing part 102and the speech synthesis section 103.

[0064]The voice recognition part 101 outputs the information which the text of the contents of utterance and others accompany as a recognition result to the sounds (a user's utterance sentence etc.) inputted from speech input systemssuch as the microphone 16.

[0065]The conversational-processing part 102 analyzes the information which the voice recognition part 101 outputsand generates a response sentence suitable based on the mobile robot's 1 growth degree.

[0066]The speech synthesis section 103 synthesizes voice to the text obtained mainly as a machinery conversation resultand outputs synthesized speech.

[0067]The functional constitution of the voice recognition part 101 is typically shown in drawing 4.

[0068]The sound uttered from the user is incorporated via the microphone 16is sampled by the AD translation part 101-1and is changed into digital data.

[0069]This digital data is changed into the parameter showing the acoustical feature of soundssuch as spectrum data for every suitable framein the feature extraction part 101-2.

[0070]And in the comparative collation part 101-3comparative collation processing with the inputted parameter is performed based on data base informationsuch as the acoustic model 101-4the dictionary 101-5and the grammar 101-6and it outputs as a speech recognition result.

[0071]Herethe acoustic model 101-4 is a model holding the acoustical featuressuch as each phoneme used in the target language (for exampleJapanese)and a syllableand Hidden Markov Model (HMM:Hidden Markov Model) etc. are used. (Hidden Markov Model expresses a time change of the feature inside a phoneme or a word in some the stateand models transition between the stateand the situation of the spectrum characteristic quantity in each state probable.) Since Hidden Markov Model deals with change of an audio signal probableit can express

fluctuation of voice inputtingetc. suitably compared with DP (Dynamic Programming) matching method.

[0072]The dictionary 101–5 is the database which described the information about the pronunciation of each word used as a recognition object.

[0073]It is the database which described how each word described by the dictionary could chain the grammar 101–6and the description based on context-free grammar (CFG)statistical word chain probability (N–gram)etc. are used.

[0074]The functional constitution of the conversational-processing part 102 is typically shown in drawing 5.

[0075]If a speech recognition result is obtained by the voice recognition part 101using the dictionary 102–3 and the grammar 102–4 for analysisthe language processing part 102–1 will perform analysessuch as a morphological analysis and syntax analysisto a recognition resultand will extract language information which constitutes a speech recognition resultsuch as information on a wordand syntactic information.

[0076]The language processing part 102–1 extracts the meaning of the voice utterance inputted by the user based on the contents described by the dictionaryan intentionetc. A speech recognition result may be given as a text of natural languageand may be given by a character string including two or more recognition candidates or graph structure.

[0077]The conversation generation part 102–2 generates a suitable response sentence with reference to the conversation production rule 102–5 based on the meaning information etc. which were extracted from utterance.

[0078]Herethe dictionary 102–3 holds informationincluding part-of-speech information required in order to apply the notation and the grammar for analysis of each word etc.and the individual meaning information of a word.

[0079]The grammar 102–4 for analysis can analyze the text which is what described the restrictions about a word chainand was inputted based on it based on the information on each word described by the dictionary 102–3.

Specificallyregular grammarcontext-free grammarstatistical word chain probabilityetc. are used. When it includes even semantic analysisthe grammar based on the language theory included semanticssuch as HPSGis used.

[0080]The conversation production rule 102–5 is the database which described the rule what kind of response sentence to generate how.

[0081]The functional constitution of the speech synthesis section 103 is typically shown in drawing 6.

[0082]If a text (namelyking head generated by the conversational-processing part 102) to synthesize voice from is givenFirstusing the dictionary 103–4 and the grammar 103–5 for analysisthe text analyzing part 103–1 performs analysessuch as a morphological analysis and syntax analysisto an input textand extracts the information which is needed by latter synthesis by rule.

[0083]Hereas information which is needed by synthesis by rulethe information about the position of a pausethe information which controls an accent and intonationand the information about the pronunciation of each word are mentioned.

- [0084]Based on these informationthe synthesis-by-rule part 103–2 connects phoneme piece datagives prosody informationsuch as a pausean accentand intonationappropriatelyand generates the digital data of an applicable sound.
- [0085]And the DA translation part 103–3 changes the voice data of digital format into analog dataand performs voice response.
- [0086]The dictionary 103–4 is a database holding the information about the part-of-speech information and reading of each wordand an accent.
- [0087]Based on the information on each word described by the dictionary 103–4the grammar 103–5 for analysis is what described the restrictions about a word chainand can analyze an input text based on it.
- [0088]The phoneme piece data 103–6 is a database holding the data of the element of a sound connected regularly. As an element of a soundit holdsfor example in the form of valve flow coefficient chain (consonant vowel chain)VCV concatenationa CVC chainetc.
- [0089]Subsequentlythe dialog mechanism of the voice base concerning this invention is explained.
- [0090]Belowthe growth level in the mobile robot's 1 time t is set to phi (t) and (0 <=phi(t) <=100)a set of the word registered into the dictionary 102–3 is set to W (w\*\*W)and a set of the syntax registered into the grammar 102–4 for analysis is set to C (c\*\*C). Hereintelligence level I (w) and I (c) are matched with each element of W and C. I (w) shows the intelligence level of the word "w"and I (c) shows the intelligence level of the syntax c.
- [0091]An intelligence level evaluates and expresses intelligence required for the mobile robot 1 to understand a word and syntax. For examplebased on the following standardsan intelligence level can be set up in 0–100.
- [0092](1) What has a simpler semantic content which a word shows has a lower intelligence leveland a more complicated thing has a higher intelligence level.
- (2) An intelligence level is so low that the subject which a word means can understand easily to the mobile robot 1. (The ex. mobile robot's 1 hand and foot have a low intelligence level.)
- (3) What requires the recognition capability by advanced vision processing has a high intelligence level.
- (4) Change an intelligence level about syntax according to the complexity.
- [0093]While a certain conversational sentences s are inputted and word w<sub>1</sub>w<sub>2</sub>--w<sub>hw</sub> are recognized as a word in the language processing part 102–1Syntax c<sub>1</sub>c<sub>2</sub>--when c<sub>nc</sub> has been recognized as a grammar analysis resultit can ask for intelligence level [ of the input conversational sentences s ] I (s)for example with a following formula.
- [0094]
- [Equation 1]

[0095]Howeverin an upper typek<sub>w</sub> and k<sub>o</sub> are the weighting factors of a word and syntaxrespectively.

[0096]In the conversation generation part 102–2a response sentence is generated using growing–degrees [ of meaning information extracted from utteranceand the mobile robot 1 ] phi (t)and intelligence level [ of the conversational sentences s ] I (s).

[0097]When response  $r_i(s)$  to the recognized conversational sentences sand ( $i=0-nr$ ) are registered one or morethe intelligence level  $I(r_i(s))$  is set up according to complexity of the contents of a response. This setting out may be matched with the response sentence itselfand may be calculated from a word and syntax which constitute a response sentence using the above–mentioned formula [one number]. About a meaning of a valueit changes with two kinds of methods explained below.  $r_0(s)$  considers it as the contents of a response which show that he was not able to understand an input sentence of “not understanding”and makes an intelligence level the minimum value  $I(r_0(s)) = 0$ .

[0098]Belowtwo kinds of methods of choosing a response according to growth of the mobile robot 1 using growing–degrees [ of the three parameters 1 about intelligencei.e.a mobile robot] phi (t) and intelligence level [ of the text s ] I and intelligence level  $I(r_i(s))$  of (s) and a response sentence are explained.

[0099][A method 1] How to choose according to relative–value (degree of transcendency)  $H(ts) = \phi(t) - I(s)$  of growing–degrees [ of the mobile robot 1 ] phi (t) to intelligence level [ of the text s ] I (s) is explained first.

[0100]Hereherea relative value of growing–degrees [ of the mobile robot 1 ] phi (t) to intelligence level [ of the input sentence s ] I (s) is defined as the degree H of transcendency (ts)and distance of  $H(ts)$  and  $I(r_i(s))$  chooses the nearest thing. That isresponse  $r_i(s)$  which fills the following formulas is chosen.

[0101]

[Equation 2]

[0102]According to this methodthe mobile robot's 1 growing degrees phi (t) from intelligence level [ of the time s of  $H(ts) \leq 0$ i.e.an input sentence] I (s) when lowSince the response which is  $\min(I(r_i(s))) = I(r_0(s))$  of “not understanding” is certainly chosenthe effect that the mobile robot 1 understands a text with a high intelligence level according to his growing degrees phi (t) can be given. The range of the preset value of intelligence level  $I(r_i(s))$  of a response sentence is set to  $0 \leq I(r_i(s)) \leq 100 - I(s)$ and it depends for it on intelligence level [ of the input sentence s ] I (s).

[0103]In a method mentioned aboveaccording to a value of growing–degrees [ of the mobile robot 1 ] phi (t)a response becomes settled uniquelyand a response may become monotonous. Thenby asking for the probability–of–occurrence density  $p(r_i(s))$  of response  $r_i(s)$  to the growing degrees phi (t) instead of  $\min|I(r_i(s)) - H(ts)|$  using the following formulasThe probability of occurrence  $P(r_i(s))$  of response  $r_i(s)$  can be searched for. Howeversigma is a variance showing a size of Gaussian distribution.

[0104]

[Equation 3]

[0105]

[Equation 4]

[0106]If the probability of occurrence  $P(r_i(s))$  and the random number generator of each response  $r_i$  are combinedthe response that the probability that a complicated response will occur according to the mobile robot's 1 growing degrees  $\phi(t)$  becomes high is generable. In this caseif the value of the intelligence level prepares many near responseseven if it inputs the same conversational sentences as the mobile robot 1 of the same growing degreesvarious responses according to probability will come on the contraryand can generate a more natural response.

[0107][A method 2] It ranks second and how to choose a response sentence by intelligence level  $I(r_i(s))$  of response sentence  $r_i(s)$  and the mobile robot's 1 growing degrees  $\phi(t)$  is independently explained to intelligence level [ of the text s ]  $I(s)$ .

[0108]Here a response with intelligence level  $I(r_i(s))$  nearest to the mobile robot's 1 growing degrees  $\phi(t)$  is chosen as below according to a formula.

[0109]

[Equation 5]

[0110]In this methodsince growing-degrees [ of the mobile robot 1 ]  $\phi(t)$  will be directly compared with the intelligence level of a response sentenceit is not necessary to calculate intelligence level [ of the input sentence s ]  $I(s)$ . The range of the preset value of intelligence level  $I(r_i(s))$  of response sentence  $r_i(s)$  is  $0 <= I(r_i(s)) <= 100$ .

[0111]Also in this methodif the following formulas are usedthe probability of occurrence of response sentence  $r_i(s)$  can be searched for.

[0112]

[Equation 6]

[0113]

[Equation 7]

[0114]The example which applied the method of using probability density in the [method 1] mentioned above at the end is explained.

[0115]Voice input of the conversational sentences s "a tomato is liking ?" is carried outand they assume that each parameter and an intelligence level were set up as follows.

[0116]  
[Equation 8]

[0117]The conversation generation part 102-2 generates the candidate of the response sentence to this input sentence s. The contents and the intelligence level of each response sentence which were generated are illustrated below.

[0118]  
[Equation 9]

[0119]Each response sentence  $r_0$  which calculated them in such setting out by having changed the mobile robot's 1 growing degrees within the limits of  $0 \leq \phi(t) \leq 100$  (s)A graph which plotted the probability density  $p(r_0(s))$  $p(r_1(s))$ and  $p(r_2(s))$  of  $r(s)$  and ,and  $r_2(s)$  is shown in drawing 7. A graph which plotted the probability  $P(r_0(s))$  $P(r_1(s))$ and  $P(r_2(s))$  of each response sentence  $r(s)$  and  $r_0(s)$  and ,and  $r_2(s)$  is shown in drawing 8.

[0120]Response  $r_0$ [ of contents that he cannot understand the input sentence s ] (s) = "a tomato is not known." occurs in high probability until the mobile robot's 1 growing degrees  $\phi(t)$  approach intelligence level [ of the input sentence s ] I(s) so that it may understand with reference to drawing 8.

[0121]And the probability of occurrence of response sentence  $r_0(s)$  fallsand the probability of occurrence of response sentence  $r_1$ [ of the following growth step ] (s) = "he likes a tomato." increases as the mobile robot 1 grows with progress of the time t. The probability of occurrence of response sentence  $r_2$ [ of the following growth step ] (s) = "although he likes a tomatoit dislikes tomato juice." becomes high as the probability of occurrence of response sentence  $r_1(s)$  will fallif it furthermore grows up.

[0122]Thusby choosing a response according to probability densityit is expected that probability of generating an intellectual response sentence according to the mobile robot's 1 growing degrees  $\phi(t)$  will become high. Even if it combines and inputs the same conversational sentences as the mobile robot 1 of the same growing degreesvarious responses according to probability will come on the contraryand can generate a more natural response.

[0123][Supplement] It has explained in detail about this inventionreferring to a specific example above. Howeverit is obvious that a person skilled in the art can accomplish correction and substitution of this example in the range which does not deviate from a gist of this invention.

[0124]Although this specification has mainly explained an embodiment which applied this invention to a <sub>pet</sub> <sub>robot</sub> a gist of this invention is not necessarily limited to a product called a "robot." That isif it is a mechanical apparatus which performs movement modeled on operation of human being using an electric or magnetic operationeven if it is a product which belongsfor example to other fields of industrysuch as a toythis invention is applicable similarly. Of coursethis invention

is applicable also to animation using a character which operates with computer graphics (CG)etc.

[0125]In shortwith a gestalt of illustrationthis invention has been indicated and it should not be interpreted restrictively. In order to judge a gist of this inventiona column of a claim indicated at the beginning should be taken into consideration.

[0126]

[Effect of the Invention]As a full account was given aboveaccording to this inventiona great leg formula mobile robot with interactive service with a user and a method for controlling the same can be provided.

[0127]According to this inventiona great leg formula mobile robot which can perform the dialog according to change of the internal stateand a method for controlling the same can be provided.

[0128]According to this inventiona great leg formula mobile robot which can change the vocabulary used according to the degree of growth of an internal statethe pattern of a responseetc.and can grow up interactive serviceand a method for controlling the same can be provided.

[0129]As for the leg formula mobile robot concerning this inventiona dialog grows according to the degree of growth of an internal state. For examplethe vocabulary to be used and the pattern of a response change. Although only an easy sentence speaks the probability of response generating set up a priori at first by making it change according to growth and he does not understanda complicated response can be gradually performed to a complicated sentence. Although only the name of the portion of the body is memorized at first by combining such grown up type interactive service with a learning functiona thing unrelated to the bodiessuch as a name etc. of the object currently seen graduallyalso comes to be memorized.

[0130]According to this inventionthe growth possibility of a robot can be expressed by changing the contents of the dialog according to the growth degree of a robot. When the contents of the dialog changefreshness can be given to a dialog with a user and interactive service for which a user is not made to sense weariness can be realized.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1]It is a figure showing the appearance composition of the mobile robot 1 which performs the leg formula walk by the limbs with which operation is presented in this invention.

[Drawing 2]It is a figure showing typically the hardware constitutions of the control unit which controls operation of the mobile robot 1.

[Drawing 3]It is a figure showing typically the functional constitution of the dialog mechanism of a voice base realized by the program execution by the control system 100.

[Drawing 4]It is a figure showing the functional constitution of the voice

recognition part 101 typically.

[Drawing 5]It is a figure showing the functional constitution of the conversational-processing part 102 typically.

[Drawing 6]It is a figure showing the functional constitution of the speech synthesis section 103 typically.

[Drawing 7]Each response sentence  $r_0$  which calculated them by having changed the mobile robot's 1 growing degrees within the limits of  $0 \leq \phi(t) \leq 100$  (s)It is the graph which plotted the probability density  $p(r_0(s))p(r_1(s))$ and  $p(r_2(s))$  of  $r(s)$  and  $r_1$ and  $r_2(s)$ .

[Drawing 8]Each response sentence  $r_0$  which calculated them by having changed the mobile robot's 1 growing degrees within the limits of  $0 \leq \phi(t) \leq 100$  (s)It is the graph which plotted the probability  $P(r_0(s))P(r_1(s))$ and  $P(r_2(s))$  of  $r(s)$  and  $r_1$ and  $r_2(s)$ .

[Description of Notations]

1 -- Mobile robot

2 -- Idiosoma unit

3 -- Head unit

4 -- Tail

6A-6D -- Leg unit

7 -- Head joint

8 -- Tail joint

9A-9D -- Femoral region unit

10A-10D -- Leg part unit

11A-11D -- Hip joint

12A-12D -- Knee joint

15 -- Camera (image recognition section)

16 -- Microphone

17 -- Loudspeaker

18 -- Touch sensor

19 -- LED indicator

20 -- Remote controller receive section

21 -- Battery

23 -- Actuator

30 -- Command reception part

31 -- External sensor

33 -- Battery sensor

34 -- Heat sensor

35 -- Internal sensor

100 -- Control system

101 -- Voice recognition part

102 -- Conversational-processing part

103 -- Speech synthesis section

---